ABSTRACT

The present invention relates generally to the field of video-camera systems, such as a video conferencing systems, and more particularly to video camera targeting systems that locate and acquire targets using an input characterizing a target and a machine-classification system to assist in target acquisition responsively to that input. In some embodiments, the characterization and classification are employed together with one or more inputs of other modalities such as gesture-control. In one example of the system in operation, an operator is able to make pointing gestures toward an object and, simultaneously speak a sentence identifying the object to which the speaker is pointing. At least one term of the sentence, presumably, is associated with a machine-sensible characteristic by which the object can be identified. The system captures and processes the voice and gesture inputs and re-positions a PTZ video camera to focus on the object that best matches both the characteristics and the gesture. Thus, the PTZ camera is aimed based upon the inputs the system receives and the system's ability to locate the target by its sensors.

WP-5087.1